

RAW SEQUENCE LISTING

EFS

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/578,613
Source: 1 FWO
Date Processed by STIC: 3/14/07

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 03/14/2007

PATENT APPLICATION: US/10/578,613

TIME: 14:57:33

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt

Output Set: N:\CRF4\03142007\J578613.raw

```

3. <110> APPLICANT: Tohata, Masatoshi
4     Sawada, Kazuhisa
5     Ozaki, Katsuya
6     Kobayashi, Kazuo
7     Ogasawara, Naotake
9 <120> TITLE OF INVENTION: RECOMBINANT MICROORGANISM
11 <130> FILE REFERENCE: 289779US0PCT
13 <140> CURRENT APPLICATION NUMBER: 10/578,613
14 <141> CURRENT FILING DATE: 2006-05-08
16 <150> PRIOR APPLICATION NUMBER: PCT/JP04/16891
17 <151> PRIOR FILING DATE: 2004-11-05
19 <150> PRIOR APPLICATION NUMBER: JP 2003-379167
20 <151> PRIOR FILING DATE: 2003-11-07
22 <160> NUMBER OF SEQ ID NOS: 122
24 <170> SOFTWARE: PatentIn version 3.3
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 3150
28 <212> TYPE: DNA
29 <213> ORGANISM: Bacillus sp. KSM-S237
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (573)..(3044)
36 <220> FEATURE:
37 <221> NAME/KEY: sig_peptide
38 <222> LOCATION: (573)..(659)
40 <400> SEQUENCE: 1
41 gatttgccga tgcaacaggc ttatatattag aggaaatttc tttttaaaatt gaatacggaa      60
43 taaaatcagg taaacaggtc ctgattttat ttttttgagt tttttagaga actgaagatt      120
45 gaaataaaaag tagaagacaa aggacataag aaaattgcat tagttttaat tatagaaaac      180
47 gcctttttat aattatttat acctagaacg aaaatactgt ttcgaaagcg gtttactata      240
49 aaaccttata ttccggctct tttttaaaac agggggtaaa aattcactct agtattctaa      300
51 tttcaacatg ctataataaa tttgtaagac gcaatatgca tctctttttt tacgatatat      360
53 gtaagcgggtt aaccttggtgc tatatgccga ttttaggaagg ggggtagatt gagtcaagta      420
55 gtaataatat agataactta taagttgttg agaagcagga gagcatctgg gttactcaca      480
57 agttttttta aaactttaac gaaagcactt tcggtaatgc ttatgaattt agctatttga      540
59 ttcaattact ttaaaaatat ttaggaggtta at atg atg tta aga aag aaa aca      593
60                                     Met Met Leu Arg Lys Lys Thr
61                                     1                               5
63 aag cag ttg att tct tcc att ctt att tta gtt tta ctt cta tct tta      641
64 Lys Gln Leu Ile Ser Ser Ile Leu Ile Leu Val Leu Leu Leu Ser Leu
65      10                               15                               20
67 ttt ccg gca gct ctt gca gca gaa gga aac act cgt gaa gac aat ttt      689
68 Phe Pro Ala Ala Leu Ala Ala Glu Gly Asn Thr Arg Glu Asp Asn Phe

```

RAW SEQUENCE LISTING

DATE: 03/14/2007

PATENT APPLICATION: US/10/578,613

TIME: 14:57:33

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt

Output Set: N:\CRF4\03142007\J578613.raw

69	25	30	35	
71	aaa cat tta tta ggt aat gac aat gtt aaa cgc cct tct gag gct ggc	737		
72	Lys His Leu Leu Gly Asn Asp Asn Val Lys Arg Pro Ser Glu Ala Gly			
73	40 45 50 55			
75	gca tta caa tta caa gaa gtc gat gga caa atg aca tta gta gat caa	785		
76	Ala Leu Gln Leu Gln Glu Val Asp Gly Gln Met Thr Leu Val Asp Gln			
77	60 65 70			
79	cat gga gaa aaa att caa tta cgt gga atg agt aca cac gga tta cag	833		
80	His Gly Glu Lys Ile Gln Leu Arg Gly Met Ser Thr His Gly Leu Gln			
81	75 80 85			
83	tgg ttt cct gag atc ttg aat gat aac gca tac aaa gct ctt tct aac	881		
84	Trp Phe Pro Glu Ile Leu Asn Asp Asn Ala Tyr Lys Ala Leu Ser Asn			
85	90 95 100			
87	gat tgg gat tcc aat atg att cgt ctt gct atg tat gta ggt gaa aat	929		
88	Asp Trp Asp Ser Asn Met Ile Arg Leu Ala Met Tyr Val Gly Glu Asn			
89	105 110 115			
91	ggg tac gct aca aac cct gag tta atc aaa caa aga gtg att gat gga	977		
92	Gly Tyr Ala Thr Asn Pro Glu Leu Ile Lys Gln Arg Val Ile Asp Gly			
93	120 125 130 135			
95	att gag tta gcg att gaa aat gac atg tat gtt att gtt gac tgg cat	1025		
96	Ile Glu Leu Ala Ile Glu Asn Asp Met Tyr Val Ile Val Asp Trp His			
97	140 145 150			
99	gtt cat gcg cca ggt gat cct aga gat cct gtt tat gca ggt gct aaa	1073		
100	Val His Ala Pro Gly Asp Pro Arg Asp Pro Val Tyr Ala Gly Ala Lys			
101	155 160 165			
103	gat ttc ttt aga gaa att gca gct tta tac cct aat aat cca cac att	1121		
104	Asp Phe Phe Arg Glu Ile Ala Ala Leu Tyr Pro Asn Asn Pro His Ile			
105	170 175 180			
107	att tat gag tta gcg aat gag ccg agt agt aat aat aat ggt gga gca	1169		
108	Ile Tyr Glu Leu Ala Asn Glu Pro Ser Ser Asn Asn Asn Gly Gly Ala			
109	185 190 195			
111	ggg att ccg aat aac gaa gaa ggt tgg aaa gcg gta aaa gaa tat gct	1217		
112	Gly Ile Pro Asn Asn Glu Gly Trp Lys Ala Val Lys Glu Tyr Ala			
113	200 205 210 215			
115	gat cca att gta gaa atg tta cgt aaa agc ggt aat gca gat gac aac	1265		
116	Asp Pro Ile Val Glu Met Leu Arg Lys Ser Gly Asn Ala Asp Asp Asn			
117	220 225 230			
119	att atc att gtt ggt agt cca aac tgg agt cag cgt ccg gac tta gca	1313		
120	Ile Ile Ile Val Gly Ser Pro Asn Trp Ser Gln Arg Pro Asp Leu Ala			
121	235 240 245			
123	gct gat aat cca att gat gat cac cat aca atg tat act gtt cac ttc	1361		
124	Ala Asp Asn Pro Ile Asp Asp His His Thr Met Tyr Thr Val His Phe			
125	250 255 260			
127	tac act ggt tca cat gct gct tca act gaa agc tat ccg tct gaa act	1409		
128	Tyr Thr Gly Ser His Ala Ala Ser Thr Glu Ser Tyr Pro Ser Glu Thr			
129	265 270 275			
131	cct aac tct gaa aga gga aac gta atg agt aac act cgt tat gcg tta	1457		
132	Pro Asn Ser Glu Arg Gly Asn Val Met Ser Asn Thr Arg Tyr Ala Leu			
133	280 285 290 295			

RAW SEQUENCE LISTING

DATE: 03/14/2007

PATENT APPLICATION: US/10/578,613

TIME: 14:57:33

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt

Output Set: N:\CRF4\03142007\J578613.raw

135	gaa aac gga gta gcg gta ttt gca aca gag tgg gga acg agt caa gct	1505
136	Glu Asn Gly Val Ala Val Phe Ala Thr Glu Trp Gly Thr Ser Gln Ala	
137	300 305 310	
139	agt gga gac ggt ggt cct tac ttt gat gaa gca gat gta tgg att gaa	1553
140	Ser Gly Asp Gly Gly Pro Tyr Phe Asp Glu Ala Asp Val Trp Ile Glu	
141	315 320 325	
143	ttt tta aat gaa aac aac att agc tgg gct aac tgg tct tta acg aat	1601
144	Phe Leu Asn Glu Asn Asn Ile Ser Trp Ala Asn Trp Ser Leu Thr Asn	
145	330 335 340	
147	aaa aat gaa gta tct ggt gca ttt aca cca ttc gag tta ggt aag tct	1649
148	Lys Asn Glu Val Ser Gly Ala Phe Thr Pro Phe Glu Leu Gly Lys Ser	
149	345 350 355	
151	aac gca acc aat ctt gac cca ggt cca gat cat gtg tgg gca cca gaa	1697
152	Asn Ala Thr Asn Leu Asp Pro Gly Pro Asp His Val Trp Ala Pro Glu	
153	360 365 370 375	
155	gaa tta agt ctt tct gga gaa tat gta cgt gct cgt att aaa ggt gtg	1745
156	Glu Leu Ser Leu Ser Gly Glu Tyr Val Arg Ala Arg Ile Lys Gly Val	
157	380 385 390	
159	aac tat gag cca atc gac cgt aca aaa tac acg aaa gta ctt tgg gac	1793
160	Asn Tyr Glu Pro Ile Asp Arg Thr Lys Tyr Thr Lys Val Leu Trp Asp	
161	395 400 405	
163	ttt aat gat gga acg aag caa gga ttt gga gtg aat tcg gat tct cca	1841
164	Phe Asn Asp Gly Thr Lys Gln Gly Phe Gly Val Asn Ser Asp Ser Pro	
165	410 415 420	
167	aat aaa gaa ctt att gca gtt gat aat gaa aac aac act ttg aaa gtt	1889
168	Asn Lys Glu Leu Ile Ala Val Asp Asn Glu Asn Asn Thr Leu Lys Val	
169	425 430 435	
171	tcg gga tta gat gta agt aac gat gtt tca gat ggc aac ttc tgg gct	1937
172	Ser Gly Leu Asp Val Ser Asn Asp Val Ser Asp Gly Asn Phe Trp Ala	
173	440 445 450 455	
175	aat gct cgt ctt tct gcc aac ggt tgg gga aaa agt gtt gat att tta	1985
176	Asn Ala Arg Leu Ser Ala Asn Gly Trp Gly Lys Ser Val Asp Ile Leu	
177	460 465 470	
179	ggt gct gag aag ctt aca atg gat gtt att gtt gat gaa cca acg acg	2033
180	Gly Ala Glu Lys Leu Thr Met Asp Val Ile Val Asp Glu Pro Thr Thr	
181	475 480 485	
183	gta gct att gcg gcg att cca caa agt agt aaa agt gga tgg gca aat	2081
184	Val Ala Ile Ala Ala Ile Pro Gln Ser Ser Lys Ser Gly Trp Ala Asn	
185	490 495 500	
187	cca gag cgt gct gtt cga gtg aac gcg gaa gat ttt gtc cag caa acg	2129
188	Pro Glu Arg Ala Val Arg Val Asn Ala Glu Asp Phe Val Gln Gln Thr	
189	505 510 515	
191	gac ggt aag tat aaa gct gga tta aca att aca gga gaa gat gct cct	2177
192	Asp Gly Lys Tyr Lys Ala Gly Leu Thr Ile Thr Gly Glu Asp Ala Pro	
193	520 525 530 535	
195	aac cta aaa aat atc gct ttt cat gaa gaa gat aac aat atg aac aac	2225
196	Asn Leu Lys Asn Ile Ala Phe His Glu Glu Asp Asn Asn Met Asn Asn	
197	540 545 550	
199	atc att ctg ttc gtg gga act gat gca gct gac gtt att tac tta gat	2273

RAW SEQUENCE LISTING

DATE: 03/14/2007

PATENT APPLICATION: US/10/578,613

TIME: 14:57:33

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt

Output Set: N:\CRF4\03142007\J578613.raw

200	Ile	Ile	Leu	Phe	Val	Gly	Thr	Asp	Ala	Ala	Asp	Val	Ile	Tyr	Leu	Asp	
201				555					560					565			
203	aac	att	aaa	gta	att	gga	aca	gaa	gtt	gaa	att	cca	gtt	gtt	cat	gat	2321
204	Asn	Ile	Lys	Val	Ile	Gly	Thr	Glu	Val	Glu	Ile	Pro	Val	Val	His	Asp	
205			570					575					580				
207	cca	aaa	gga	gaa	gct	gtt	ctt	cct	tct	gtt	ttt	gaa	gac	ggg	aca	cgt	2369
208	Pro	Lys	Gly	Glu	Ala	Val	Leu	Pro	Ser	Val	Phe	Glu	Asp	Gly	Thr	Arg	
209			585					590					595				
211	caa	ggg	tgg	gac	tgg	gct	gga	gag	tct	ggg	gtg	aaa	aca	gct	tta	aca	2417
212	Gln	Gly	Trp	Asp	Trp	Ala	Gly	Glu	Ser	Gly	Val	Lys	Thr	Ala	Leu	Thr	
213	600					605					610				615		
215	att	gaa	gaa	gca	aac	ggg	tct	aac	gcg	tta	tca	tgg	gaa	ttt	gga	tat	2465
216	Ile	Glu	Glu	Ala	Asn	Gly	Ser	Asn	Ala	Leu	Ser	Trp	Glu	Phe	Gly	Tyr	
217				620						625					630		
219	cca	gaa	gta	aaa	cct	agt	gat	aac	tgg	gca	aca	gct	cca	cgt	tta	gat	2513
220	Pro	Glu	Val	Lys	Pro	Ser	Asp	Asn	Trp	Ala	Thr	Ala	Pro	Arg	Leu	Asp	
221			635						640					645			
223	ttc	tgg	aaa	tct	gac	ttg	gtt	cgc	ggg	gag	aat	gat	tat	gta	gct	ttt	2561
224	Phe	Trp	Lys	Ser	Asp	Leu	Val	Arg	Gly	Glu	Asn	Asp	Tyr	Val	Ala	Phe	
225			650					655					660				
227	gat	ttc	tat	cta	gat	cca	gtt	cgt	gca	aca	gaa	ggc	gca	atg	aat	atc	2609
228	Asp	Phe	Tyr	Leu	Asp	Pro	Val	Arg	Ala	Thr	Glu	Gly	Ala	Met	Asn	Ile	
229			665					670					675				
231	aat	tta	gta	ttc	cag	cca	cct	act	aac	ggg	tat	tgg	gta	caa	gca	cca	2657
232	Asn	Leu	Val	Phe	Gln	Pro	Pro	Thr	Asn	Gly	Tyr	Trp	Val	Gln	Ala	Pro	
233	680					685						690				695	
235	aaa	acg	tat	acg	att	aac	ttt	gat	gaa	tta	gag	gaa	gcg	aat	caa	gta	2705
236	Lys	Thr	Tyr	Thr	Ile	Asn	Phe	Asp	Glu	Leu	Glu	Glu	Ala	Asn	Gln	Val	
237				700						705					710		
239	aat	ggg	tta	tat	cac	tat	gaa	gtg	aaa	att	aac	gta	aga	gat	att	aca	2753
240	Asn	Gly	Leu	Tyr	His	Tyr	Glu	Val	Lys	Ile	Asn	Val	Arg	Asp	Ile	Thr	
241			715						720					725			
243	aac	att	caa	gat	gac	acg	tta	cta	cgt	aac	atg	atg	atc	att	ttt	gca	2801
244	Asn	Ile	Gln	Asp	Asp	Thr	Leu	Leu	Arg	Asn	Met	Met	Ile	Ile	Phe	Ala	
245			730					735					740				
247	gat	gta	gaa	agt	gac	ttt	gca	ggg	aga	gtc	ttt	gta	gat	aat	gtt	cgt	2849
248	Asp	Val	Glu	Ser	Asp	Phe	Ala	Gly	Arg	Val	Phe	Val	Asp	Asn	Val	Arg	
249			745					750					755				
251	ttt	gag	ggg	gct	gct	act	act	gag	ccg	gtt	gaa	cca	gag	cca	gtt	gat	2897
252	Phe	Glu	Gly	Ala	Ala	Thr	Thr	Glu	Pro	Val	Glu	Pro	Glu	Pro	Val	Asp	
253	760					765					770				775		
255	cct	ggc	gaa	gag	acg	cca	cct	gtc	gat	gag	aag	gaa	gag	aaa	aaa	gaa	2945
256	Pro	Gly	Glu	Glu	Thr	Pro	Pro	Val	Asp	Glu	Lys	Glu	Ala	Lys	Lys	Glu	
257				780						785					790		
259	caa	aaa	gaa	gca	gag	aaa	gaa	gag	aaa	gaa	gca	gta	aaa	gaa	gaa	aag	2993
260	Gln	Lys	Glu	Ala	Glu	Lys	Glu	Glu	Lys	Glu	Ala	Val	Lys	Glu	Glu	Lys	
261			795						800					805			
263	aaa	gaa	gct	aaa	gaa	gaa	aag	aaa	gca	gtc	aaa	aat	gag	gct	aag	aaa	3041
264	Lys	Glu	Ala	Lys	Glu	Glu	Lys	Lys	Ala	Val	Lys	Asn	Glu	Ala	Lys	Lys	

RAW SEQUENCE LISTING

DATE: 03/14/2007

PATENT APPLICATION: US/10/578,613

TIME: 14:57:33

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt

Output Set: N:\CRF4\03142007\J578613.raw

```

265      810      815      820
267 aaa taatctatta aactagttat agggttatct aaaggtctga tgtagatctt      3094
268 Lys
271 ttagataacc tttttcttgc ataactggac acagagttgt tattaaagaa agtaag      3150
274 <210> SEQ ID NO: 2
275 <211> LENGTH: 824
276 <212> TYPE: PRT
277 <213> ORGANISM: Bacillus sp. KSM-S237
279 <400> SEQUENCE: 2
281 Met Met Leu Arg Lys Lys Thr Lys Gln Leu Ile Ser Ser Ile Leu Ile
282 1      5      10      15
285 Leu Val Leu Leu Ser Leu Phe Pro Ala Ala Leu Ala Ala Glu Gly
286      20      25      30
289 Asn Thr Arg Glu Asp Asn Phe Lys His Leu Leu Gly Asn Asp Asn Val
290      35      40      45
293 Lys Arg Pro Ser Glu Ala Gly Ala Leu Gln Leu Gln Glu Val Asp Gly
294      50      55      60
297 Gln Met Thr Leu Val Asp Gln His Gly Glu Lys Ile Gln Leu Arg Gly
298 65      70      75      80
301 Met Ser Thr His Gly Leu Gln Trp Phe Pro Glu Ile Leu Asn Asp Asn
302      85      90      95
305 Ala Tyr Lys Ala Leu Ser Asn Asp Trp Asp Ser Asn Met Ile Arg Leu
306      100     105     110
309 Ala Met Tyr Val Gly Glu Asn Gly Tyr Ala Thr Asn Pro Glu Leu Ile
310      115     120     125
313 Lys Gln Arg Val Ile Asp Gly Ile Glu Leu Ala Ile Glu Asn Asp Met
314      130     135     140
317 Tyr Val Ile Val Asp Trp His Val His Ala Pro Gly Asp Pro Arg Asp
318 145     150     155     160
321 Pro Val Tyr Ala Gly Ala Lys Asp Phe Phe Arg Glu Ile Ala Ala Leu
322      165     170     175
325 Tyr Pro Asn Asn Pro His Ile Ile Tyr Glu Leu Ala Asn Glu Pro Ser
326      180     185     190
329 Ser Asn Asn Asn Gly Gly Ala Gly Ile Pro Asn Asn Glu Glu Gly Trp
330      195     200     205
333 Lys Ala Val Lys Glu Tyr Ala Asp Pro Ile Val Glu Met Leu Arg Lys
334      210     215     220
337 Ser Gly Asn Ala Asp Asp Asn Ile Ile Ile Val Gly Ser Pro Asn Trp
338 225     230     235     240
341 Ser Gln Arg Pro Asp Leu Ala Ala Asp Asn Pro Ile Asp Asp His His
342      245     250     255
345 Thr Met Tyr Thr Val His Phe Tyr Thr Gly Ser His Ala Ala Ser Thr
346      260     265     270
349 Glu Ser Tyr Pro Ser Glu Thr Pro Asn Ser Glu Arg Gly Asn Val Met
350      275     280     285
353 Ser Asn Thr Arg Tyr Ala Leu Glu Asn Gly Val Ala Val Phe Ala Thr
354      290     295     300
357 Glu Trp Gly Thr Ser Gln Ala Ser Gly Asp Gly Gly Pro Tyr Phe Asp
358 305     310     315     320

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/578,613

DATE: 03/14/2007
TIME: 14:57:34

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt
Output Set: N:\CRF4\03142007\J578613.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31
Seq#:32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55
Seq#:56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79
Seq#:80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102
Seq#:103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120
Seq#:121,122

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/578,613

DATE: 03/14/2007

TIME: 14:57:34

Input Set : N:\efs\03_14_07\10578613_efs\289779US0PCTST25.txt

Output Set: N:\CRF4\03142007\J578613.raw